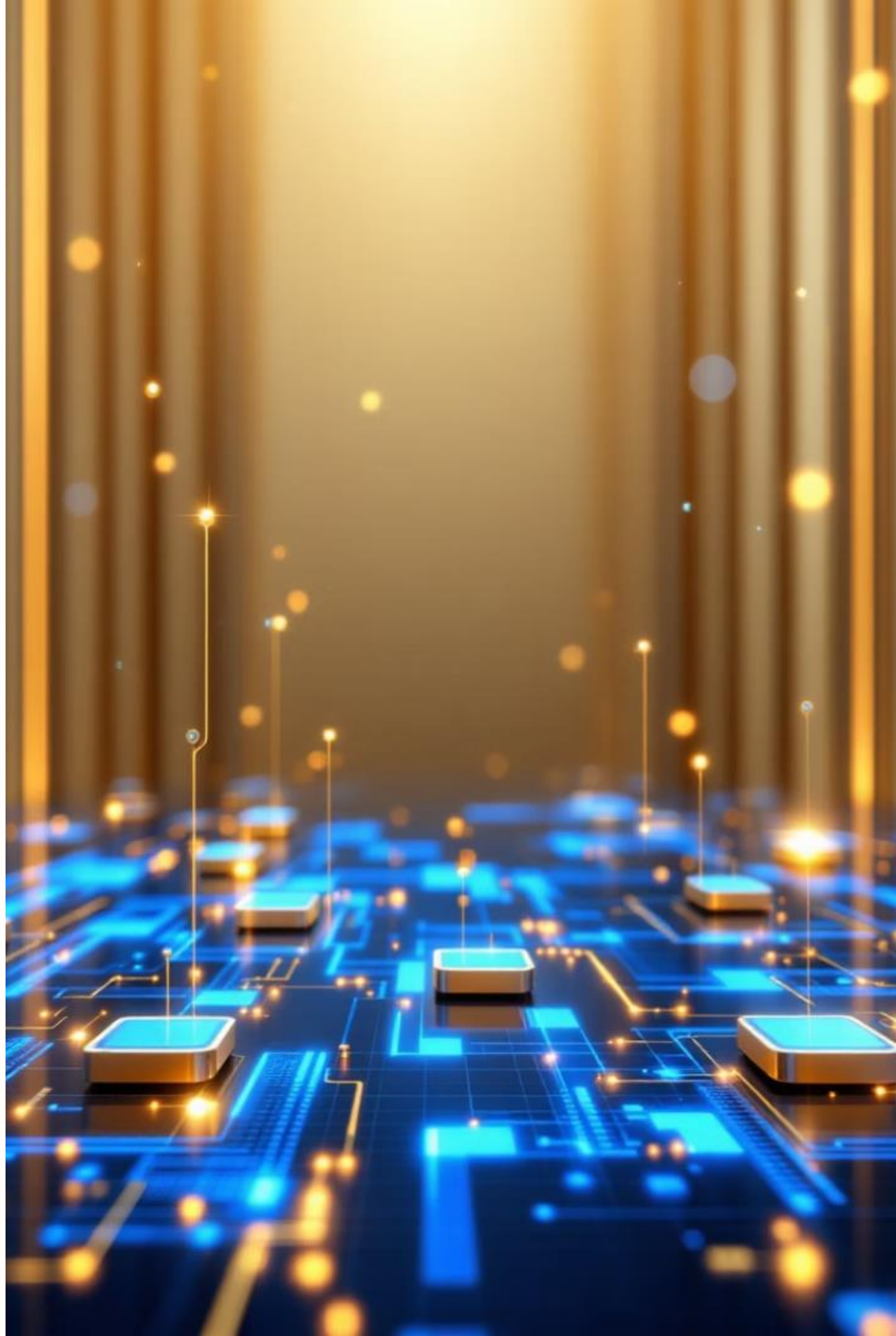
보안과 데이터:
룰 기반을 넘어, AI로 판단하는 시대

- 디지털 혁신의 중심에서 보안의 패러다임이 변화.
- 룰 기반 접근법에서 AI 기반 판단으로 전환.
- 현 시점에서 새로운 보안 관점을 제시.



보안과 데이터의 불가분 관계

감지 단계

로그와 트래픽 데이터를 수집하여 이상 징후를 감지합니다.

분석 단계

수집된 데이터를 심층적으로 분석하여 위협을 식별합니다.

대응 단계

분석 결과에 따라 적절한 방어 메커니즘을 실행합니다.

- 모든 **보안 프로세스**는 **데이터를 기반**으로 합니다.
- 효과적인 보안은 데이터의 수집, 분석, 활용에 달려 있습니다.

디지털 세계의 데이터 소우주

네트워크 데이터

트래픽 흐름, 패킷 교환, 프로토콜 사용 패턴

시스템 데이터

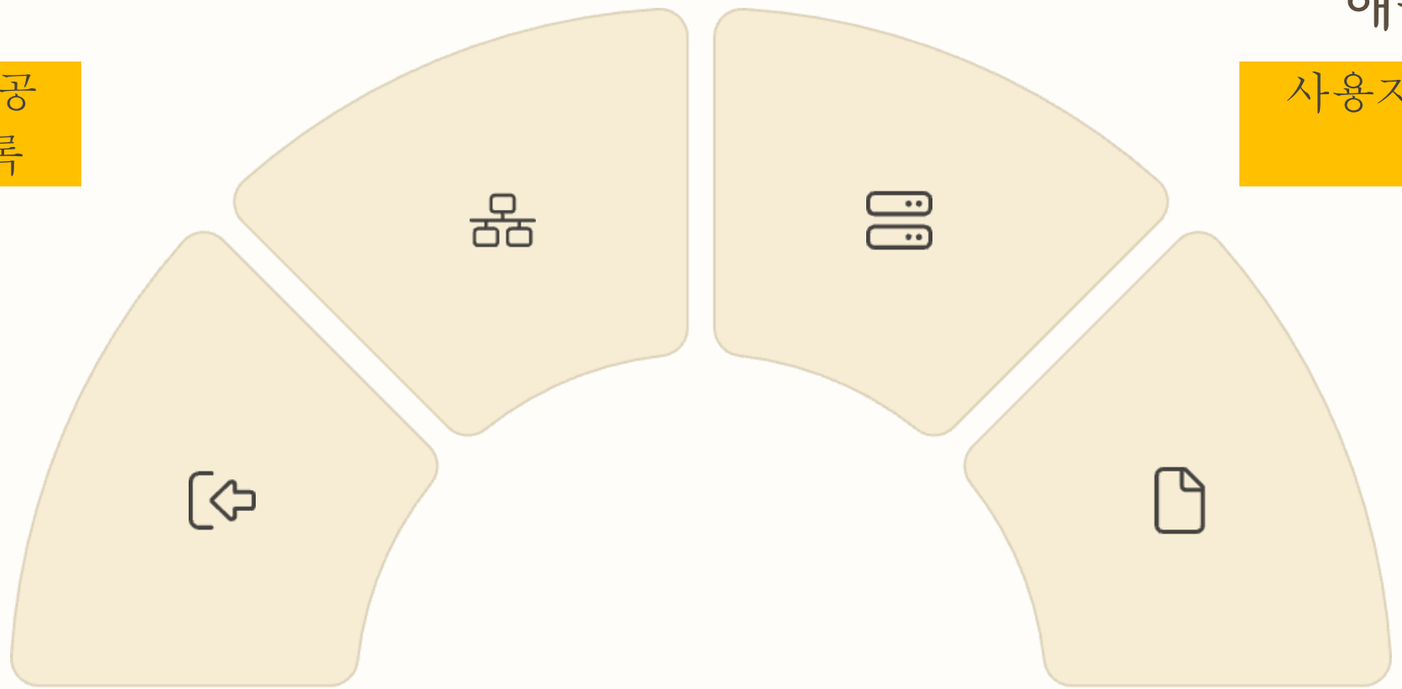
리소스 사용률, 프로세스 활동, 시스템 이벤트

인증 데이터

로그인 시도, 인증 성공/실패, 권한 변경 기록

애플리케이션 데이터

사용자 활동, 데이터 접근 패턴, API 호출



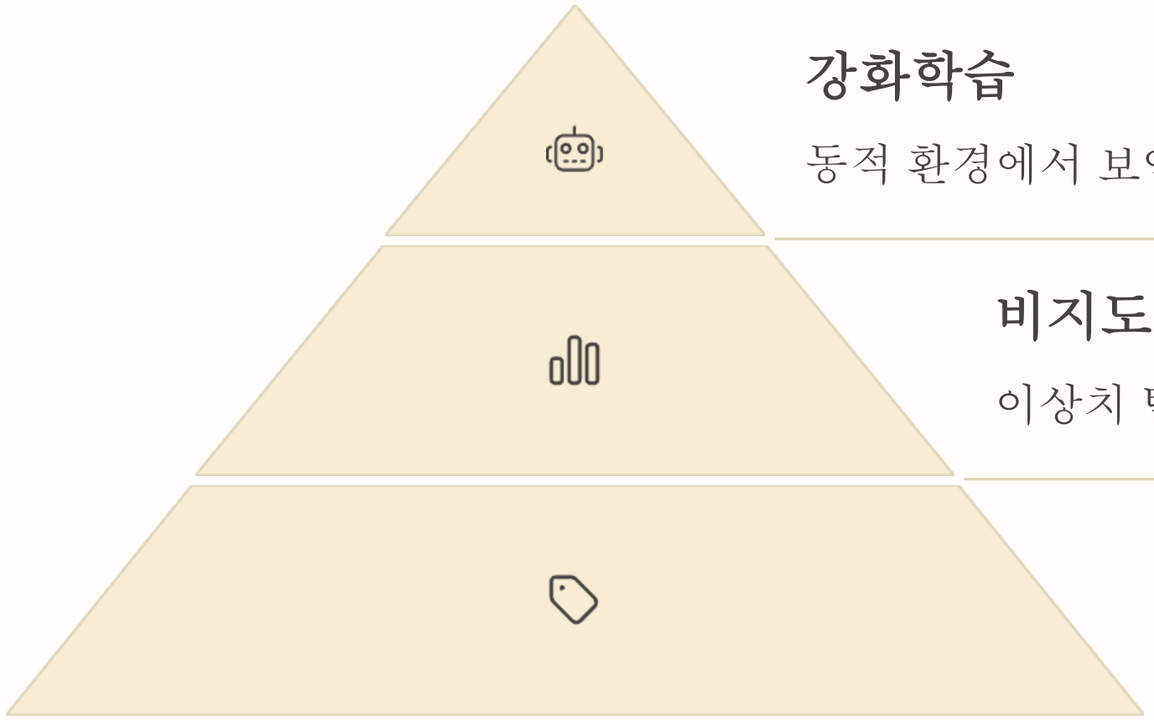
- 디지털 세계의 모든 활동은 **데이터로 기록**됩니다.
- **보안 전문가는 이 데이터를 해석하는 능력이 필수**적입니다.

보안 판단 방식의 진화

구분	룰 기반	AI 기반
설정자	보안 엔지니어	알고리즘 (자가학습)
판단 방식	조건문 (if-then)	패턴 인식 및 예측
적응성	낮음 (수동 업데이트)	높음 (자동 적응)
예시	특정 IP 차단	이상 행동 패턴 감지

- 보안 판단은 정적 룰에서 동적 AI 기반으로 진화하고 있습니다.
- 복잡한 공격에 대응하기 위한 필연적 변화입니다.

머신러닝과 보안의 융합



강화학습

동적 환경에서 보안 정책 최적화

비지도학습

이상치 탐지와 패턴 발견

지도학습

분류와 예측을 통한 위협 식별

- 머신러닝은 보안 분야에 혁명적 변화를 가져왔습니다.
- 대량의 데이터를 효과적으로 분석하여 위협을 식별합니다.

지도학습 기반 보안 판단

데이터 수집 및 라벨링

알려진 위협과 정상 행동에 대한 데이터를 수집하고 분류.

모델 학습 및 검증

머신러닝 알고리즘이 패턴을 인식하고 분류 방법을 학습.

실시간 위협 분류

학습된 모델이 새로운 데이터를 분석하여 위협 여부를 판단.

- 지도학습은 이미 알려진 패턴에 기반한 위협 탐지에 우수.
- 스팸 이메일 분류와 악성코드 탐지에 효과적.

비지도학습 기반 이상 탐지

행동 기준선 수립

시스템의 정상 상태와 행동 패턴을 학습합니다.

통계적 편차 분석

정상 범위를 벗어나는 활동을 식별합니다.

잠재적 위협 식별

비정상적인 패턴을 위협으로 분류하고 우선순위를 지정합니다.

제로데이 공격 방어

알려지지 않은 새로운 유형의 공격에도 효과적으로 대응합니다.

보안의 미래:
데이터와 AI의 시너지

99.9%

자동화 수준

보안 위협 탐지 및 대응의 자동화율

1초

반응 속도

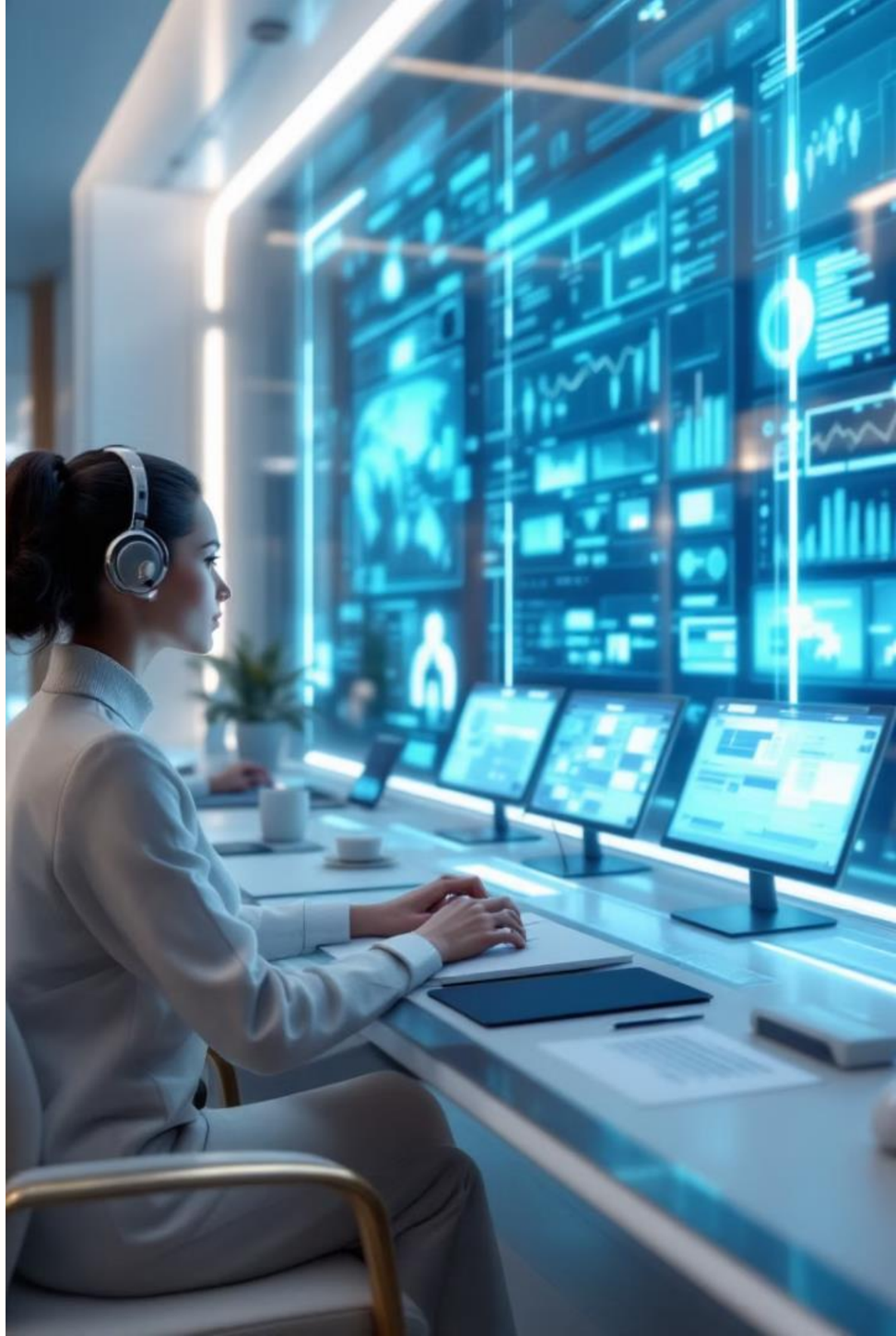
AI 기반 시스템의 위협 대응 소요 시간

10억+

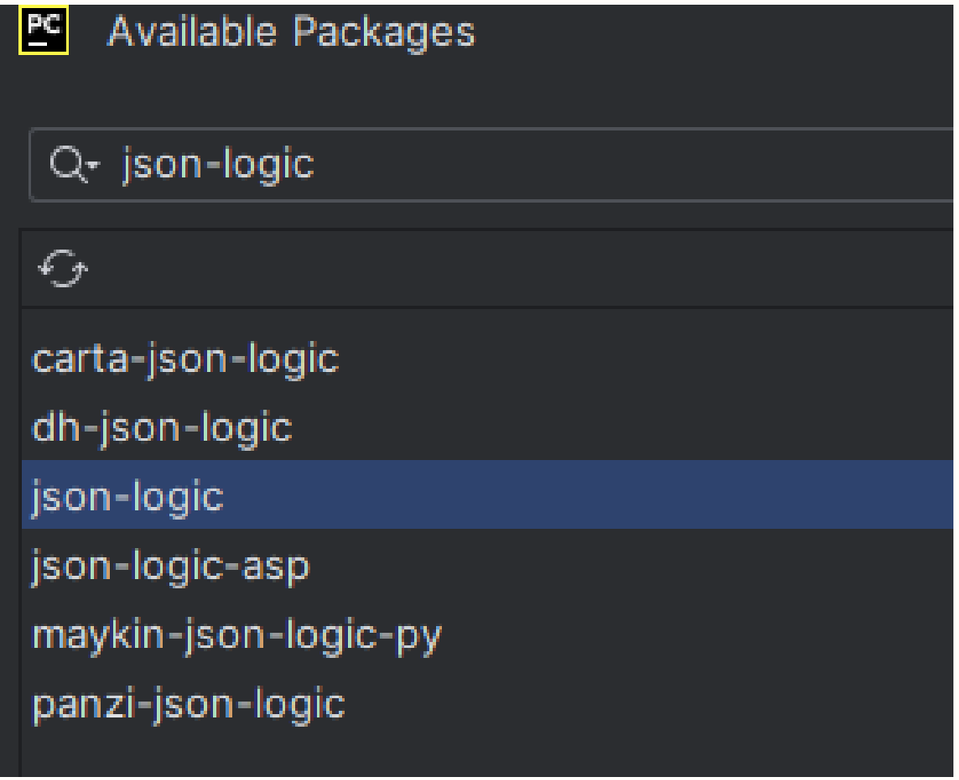
일일 데이터 포인트

AI가 분석하는 보안 이벤트 수

- **보안의 핵심은 이제 데이터와 AI**입니다.
- 현대 보안 전문가는 **데이터 분석과 AI 활용 능력이 필수**적입니다.



지도학습



- pandas
- Json-logic
- Matplotlib
- scikit-learn

C:\Users\GU\PycharmProjects\security\venv\lib\site-packages\json_logic__init__.py

op = tests.keys()[0] -> op = list(tests.keys())[0]

```python
import sys
from functools import reduce
def jsonLogic(tests, data=None):
  # You've recursed to a primitive, stop!
  if tests is None or type(tests) != dict:
    return tests

  data = data or {}

  # op = tests.keys()[0]
  op = list(tests.keys())[0]
```

op = tests.keys()[0]

op = list(tests.keys())[0]

from functools import reduce

C:\Users\GU\PycharmProjects\security\venv\lib\site-packages\json_logic__init__.py

데이터 수 부족

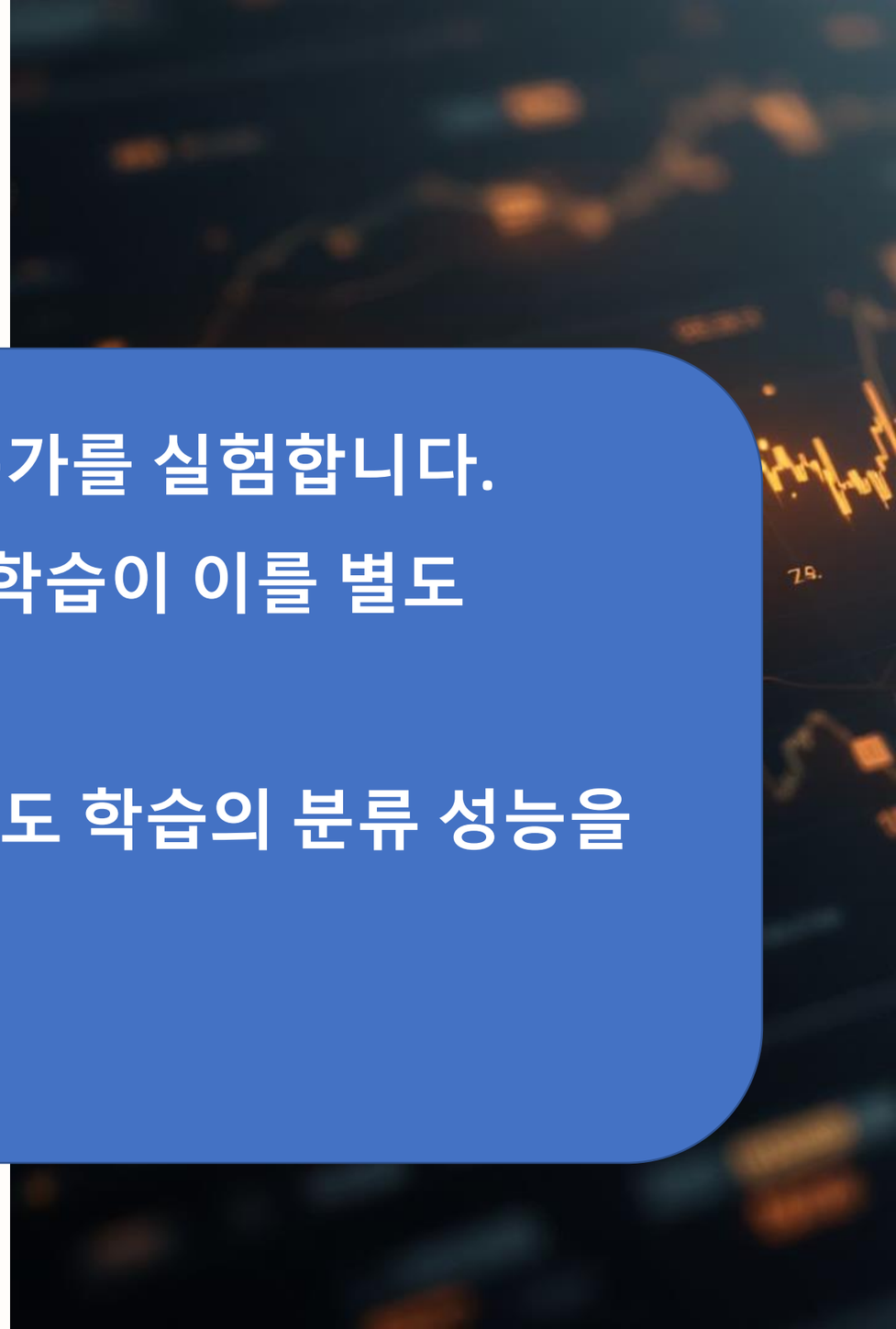
Test_mode = Const.TRAINING_BY_LOW_SAMPLE.value

데이터 특징(feature) 부족

- 과거 점수와 평균만으로 데이터 훈련
- 현재 점수만으로 데이터 훈련
- 과거 점수만으로 데이터 훈련

비 지도학습

- 군집을 기반으로 한 데이터 판단
- 이상치를 기반으로 한 데이터 판단



1. 라벨 없이도 KMeans가 군집을 나눌 수 있는가를 실험합니다.
2. 이상치 데이터 (200~300)를 추가해 비지도 학습이 이를 별도 클러스터로 인식하는지 테스트.
3. 룰 기반 결과와 군집 결과를 비교하면, 비지도 학습의 분류 성능을 시각적으로 분석할 수 있습니다.